ics Working Paper, Virginia Polytechnic Institute and State University, May 16.
Williamson, Oliver. 1963. "A Model of Rational Managerial Behavior." *A Behavioral Theory of the Firm,* edited by Richard Cyert and James March. Englewood Cliffs, N.J.: Prentice-Hall.
Wilson, James Q. 1974. "The Politics of Regulation." In *Social Responsibility and the Business Predicament,* edited by James W. McKie. Washington, D.C.: The Brookings Institution.
Wright, J. Skelly. 1978. "Commentary: Rulemaking and Judicial Review." *Administrative Law Review* 30, no. 3 (Summer): 461 f6.
Zaltm *ned*
Chc

IN

ATTACKING REGULATORY PROBLEMS
An Agenda for Research in the 1980s

Edited by
ALLEN R. FERGUSON
Public Interest Economics Foundation

BALLINGER PUBLISHING COMPANY
Cambridge, Massachusetts
A Subsidiary of Harper & Row, Publishers, Inc.

Replaces California Institute of Technology Social Science Working Paper # 298

6 EXPERIMENTAL METHODS IN POLITICAL ECONOMY
A Tool for Regulatory Research

Charles R. Plott

Regulatory research involves scholars in almost all dimensions of the social sciences. Economics, law and legal institutions, psychology, bureaucratic politics, and even voting politics become intermingled in complex ways to obscure the consequences of alternative policy options. Research undertaken to provide some insights, experience, or educated guesses necessarily confronts scholars with multidisciplinary considerations. It also involves them with situations, problems, and institutional arrangements for which there are no historical precedents, leaving them with very little shared experience to resolve scientific disagreements.

Within the last several years, both laboratory experimental methods and field experimental methods have been explored as a source of shared experiences and data for researchers and decisionmakers. The experimental technology has been advanced, and perhaps more importantly, the art

Charles Plott is professor of economics at the California Institute of Technology. Many of the references and project descriptions contained in this chapter evolved from the active participation of Paul Schoemaker, rapporteur, Dan Fulmer, John Palmisano, Laurent Ross, and Anthony Yezer in the experimental session of the conference. The efforts of Paul Schoemaker were especially important in this regard. Not only did he serve as rapporteur during the meeting, he later drafted a summary of psychological experimental methodology that helped place the methodologies developing in economics and politics in a better perspective.

of posing questions which make experimentation relevant is being developed. Scholars are beginning to see how experiences generated from simple, special case, and controlled settings can be used as criteria for determining the relative acceptability of general theories and related models of complex social systems. Scientific thinking has shifted from an almost universal belief that experimental methods are *in principle* not applicable when the questions involve politics or economics—to a qualified acknowledgment from several quarters that the methods are applicable when the problems are sufficiently simple and carefully defined. Occasionally, data from experiments systematically find their way into the policymaking process. The developments are in their infancy, and because of the structure and role of theory, they seem to differ somewhat from those that have developed within psychology.[1] However, progress is occurring as successful basic experiments provide a foundation for more complex applied work.

Essentially two different types of experiments exist. Theoretically each seems best suited to accomplish a slightly different task. The first, laboratory experimental methods, is very new to economics and politics. The capability of creating a real economic or political process within a laboratory environment was only recently acquired. The processes that can be created and sufficiently controlled are necessarily simple, relative to those that have naturally evolved. As a consequence the methods have been best suited for examining very basic and general scientific theories. However, the accumulation of basic results has led to the posing of more complicated questions, and recently some direct applications of laboratory methods to regulatory questions have been completed.

A recurrent theme in laboratory research is that models which researchers apply with great regularity to a wide range of complex processes should be expected to apply to the simple special cases of laboratory processes. If they do not, then they should be modified or rejected. Furthermore, those models which work best in simple cases have reasonable priority claim when applied to complex cases. It is by properly posing and sequencing arguments of this nature that the study of simple laboratory processes becomes relevant for the professions at large. As such, a solid empirical basis for theory is being developed.

The second class of experiments are field experiments which involve some perturbation and monitoring of a natural process. These studies (sometimes called "evaluation research") are generally more complex than laboratory experiments, and the research objectives are usually different from those of laboratory studies. Many questions cannot yet be posed within a laboratory frame of reference. Hypotheses regarding the development and evolution of institutions have not been successfully formulated for laboratory study where the institutions themselves are usually either held fixed or are treatment variables. Parameters or statistical regularities characteristic of large, diverse political and economic systems are not generally studied in laboratories. Field studies are usually designed to undertake such complex tasks and answer questions not answerable by application of any other methodology.

One major advance has involved the discovery of insights about how to properly identify research objectives so experimental methods can be usefully applied. Five different types of research objectives will be demonstrated below: (1) theory rejection, (2) theory competition, (3) measurement, (4) simulation, and (5) process design. The reader should regard these categories as an attempt to impose some structure on the literature in order to provide a coherent overview as opposed to any widely accepted methodological categorization. Any given experiment may reflect one or more of these objectives, and on occasion research objectives seem to make sense but still do not fit neatly into any of these categories.

THEORY REJECTION

Many models are applied to the behavior of complex political and economic processes. For the most part, scholars do not believe the models are "true"; instead they maintain that the models are "adequate representations" or "reasonable approximations" to the way things work. Unless such models have an internal logical inconsistency, critics will find it very difficult, if not impossible, to refute the position of an advocate who is

taking such a moderate posture. As a result, general models and theories of political and economic processes tend to die of old age, if they die at all—not because they are substantially at odds with the shared experiences of scholars. It sometimes takes decades if not more of painstaking scholarship to refute convincingly even the most ridiculous of theories.

Experiments designed to reject a theory constitute an effort to speed up the scientific sifting process. The trick involves the art of designing a *very* simple process with the essential features of the complex processes to which the target model is being applied. Models that cannot explain the behavior of simple special cases are not general models and therefore cannot be advocated as such. Furthermore, models that do not work in the simple cases, created explicitly to give a model its "best chance," can certainly be challenged as being relevant to the complex. Thus the idea is to create a situation so simple that reasonable accuracy of a model would be regarded by almost all to be a necessary condition for acceptance of the model at any level.

If the model fails to work, then the model is rejected and the experiments have been useful. On the other hand, if the model actually works in the simple case, then the experimenter has really solved no problem at all. Critics of the theory will simply dismiss the experimental results, saying that the experimental setting has nothing to do with the "real world." At best, they may be challenged to explain why the model of which they are critical might work in *any* context. On the other hand, those who support the theory will say that the experimental results simply conform to what was already "well known." [2] Thus experiments designed to reject theories may be "uninteresting" to everyone if the model intended for rejection actually works when given its "best case."

Good examples of experiments of this nature were the early market experiments conducted by Vernon Smith. By application of what has since been named induced preference theory (Plott, 1979; Smith 1976b), Smith created a very simple but very real market. One side of the market was allowed to purchase units of a commodity (called x) from the experimenter according to prespecified (marginal) cost schedules. These persons earned money by selling at a profit to those who wanted to buy the commodity. A desire to purchase the commodity was created in another group of individuals by giving them the right to redeem with the experimenter any units purchased according to a prespecified schedule of dollar amounts. Profits to individuals of this second group, which were theirs to keep, were the difference between the redemption value and the price paid by the individual. Thus for each individual, application of the theory of derived demand establishes the existence of a personal demand function for the commodity. All transactions were in cash and the magnitudes were sufficient to yield potential earnings substantially above the normal hourly rates of pay.

The market was organized as a double oral auction. Bids, offers, and contracts were made publicly on a unit-by-unit basis. The market remained open for a short period and then closed. New redemption schedules and cost schedules were distributed (for the first experiments, these remained constant) and the market opened for a second trading "day." The process continued for several trading "days."

The situation created by Smith was a very real market. It was a *simple* market relative to those found in natural environments, but it was nevertheless real and the principles that are supposed to operate in markets in general should certainly be expected to be operative there as well. The market prices, volume and income distribution should all be predictable. In particular, the basic law of supply and demand should work to precisely and predictably determine these magnitudes. If indeed economists are unable to predict the behavior of a system as simple as this one, then their ability to predict and explain the operations of much more complicated, naturally evolving markets can be questioned. The early Smith experiments were certainly posed for an attack on an important and basic theory.

The ability of the supply and demand model to predict the behavior of these simple markets is amazing. The data from four different experimental markets are shown in Figure 6–1. The market demand and supply functions are obtained from the traditional summation of the theoretical individual demand and supply functions. After only a few periods prices converge close to the equilibrium price of $2.10. Volume is close to the predicted volume of eight units, and the efficiency

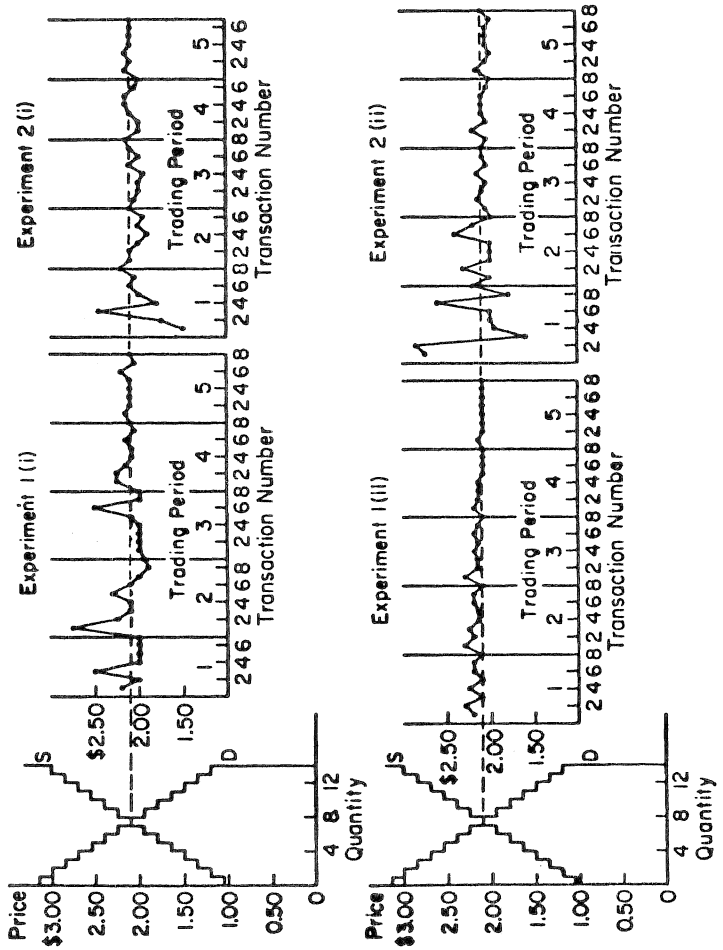


Figure 6–1 Time series of transaction prices in double oral auction markets

Source: Smith, Vernon L. 1976a. "Bidding and Auctioning Institutions: Experimental Results." In *Bidding and Auctioning for Procurement and Allocation*, edited by Yakov Amihud. New York: New York University Press. Reprinted with permission.

levels[3] of these processes are near 100 percent. The data are representative of what occurs under a variety of different, shaped curves, shifting parameters, and in markets with randomness and speculation. The model really works well.

Until very recently, academic reception to this phenomenon was cool. Economists thought that only the obvious had been demonstrated. However, interest in the behavior of these particular markets has rapidly begun to revive as scholars seek explanations of the convergence phenomenon in terms of information processing, signaling, information transfer, and the related theories of strategy. We began to realize that we really do not understand why such markets nearly always converge. An experimental design that was first constructed as a tool for theory rejection is now being used as a tool to resolve competition among theories—a function which will be discussed in the next section.

The theories do not always win when given their best chance. Psychologists (Lichtenstein and Slovic 1971, 1973; Lindman 1971; Slovic 1975; Tversky 1969) have developed a clever series of demonstrations that challenge the basic theory of preference as economists use it. Attempts by economists to modify the design or apply various economic theories to explain away the psychologists' results have failed so far (Grether, Isaac, and Plott 1979).[4] In addition to the study of pure choice behavior, a number of studies have been exploring the structure of people's risk-taking attitudes toward low-probability, high-loss hazards (natural disasters, nuclear risks, and so on). Contrary to the traditional expected utility theory (Savage 1954), people appear to be willing to take risks in the domain of low probabilities, suggesting the existence of probability thresholds (perhaps larger), below which threats are ignored or discounted. Not only are these studies finding systematic exceptions to the expected utility hypothesis (Hershey and Schoemaker, 1980; Kahneman and Tversky 1979; Schoemaker and Kunreuther, 1979), there apparently exist systematic deviations from Bayes law in the processing of information (Grether 1980).

Exactly where these observations will lead is not clear. None of the above scholars are prepared to reject the entire expected utility frame of analysis and the related use of Bayes law. No

alternative theory does so well in such a broad spectrum of areas. It is clear, however, that the theory of choice and decision, when applied at the individual level of analysis, is in need of repair. The data from these simple experiments provide a common empirical frame of reference for those who wish to explore their own ideas about possible modifications of the theory.

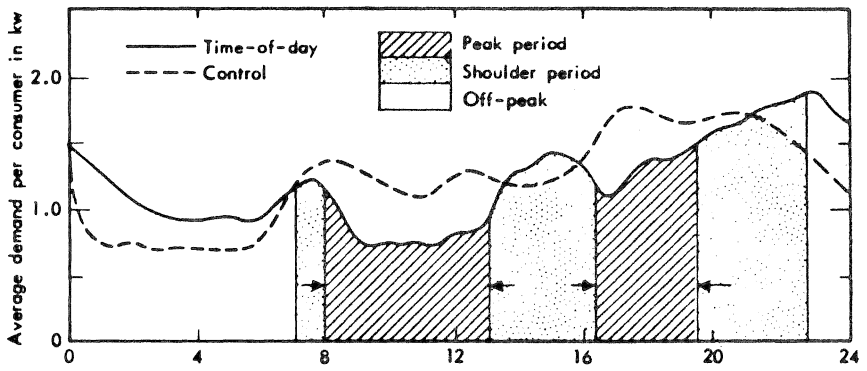
COMPETING THEORIES

As mathematical tools increasingly find applications in politics and economics, the separations among theories grows finer and finer. Even though two or more models may make drastically different predictions about what might result from a particular policy or regulation, it may be impossible to discern from historical data which is the more reliable model because the predictions of the models are usually similar. Experimental procedures are perfectly suited for this type of analysis.

Consider a controversy about the implications of peak load pricing policies for residential users (Mitchell, Manning, and Acton 1978). On one side were those who claimed that electricity users are characterized by ignorance, indifference, habit, traditions, and so on, and will not respond to a differential time-of-day tariff. Industry personnel and spokespersons with much experience and knowledge testified that their customers would not respond. Engineers with a great deal of technical expertise claimed that customers could not respond. On the other side were those who claimed that users would respond in a substituting manner typical of the optimizing postulates presupposed in many economic models. In fact the history of electricity pricing had been so uniform over the years that overwhelmingly convincing evidence could not be generated in favor of one set of ideas relative to another.

The natural thing to do was conduct a field experiment involving the actual pricing policies of a company. The results of one such experiment are given in Figure 6–2 where peak-load prices were imposed on one group while a control group faced normal prices. As can be seen, the responsiveness to time-of-day tariffs is dramatic, and the models motivated by

Figure 6–2 Time-of-day and control-group load curves in the British experiment, average winter weekday loads, residential consumers, 1969–70



Source: J.G. Boggis, personal communication. Copyright Electricity Council of London. Used by permission. Reprinted in Mitchell, Bridger M.; Willard G. Manning, Jr.; and Jan Paul Acton. 1978. *Peak-Load Pricing: European Lessons for U.S. Energy Policy.* Cambridge, Mass.: Ballinger Publishing Company.

noneconomic considerations were demonstrably less applicable. High prices at peak periods of demand shift the demand to other periods. Such results have been replicated by many other field studies and have been further refined to the specific type of rate, methods of metering, and so forth.

Results from laboratory methods have also been successful in differentiating among competing theories. For many different types of committees and circumstances the core of an appropriately defined cooperative game model can be used to predict accurately the committee decisions (Berl et al. 1976; Fiorina and Plott 1978). Models that have historically been intellectual competitors of this model are demonstrably inferior. The classic conflict between game theory and psychologically oriented theories of group processes is being placed in perspective. Within game theory itself the traditional game theoretic solution concepts that were the subject of mathematical game theory for years and remained essentially unchanged have now begun to be modified in light of experimental results (Laing and Olmsted 1978; McKelvey and Ordeshook 1978; McKelvey, Ordeshook, and Winer 1978; Plott and Rogerson 1979; Rapoport, Kahan, and Wallsten 1976).

Another example of competing ideas can be taken from market theory. Different ideas about the relative revenue-generating properties of several methods of auctioning single items have existed side by side for years. With the application of new laboratory techniques for studying auctions, the ideas are being refined and their conflicts resolved (Belovicz 1979; Smith 1967). More importantly, researchers have found observations generated by laboratory studies useful in formulating hypotheses about the properties of very large and complex markets (treasury bills) and in ascertaining which market form would generate more revenue (Tsao and Vignola 1977).

MEASUREMENT

A third objective of experimentation involves measurement. Two very different kinds of measurement exist. The first constitutes a check on the robustness of established results with slight changes in institutional parameters. The second is parameter estimation for general models.

One of the most interesting aspects of laboratory experimental methods has been the discovery that slight institutional changes in organization can have dramatic changes in the performance of the processes. The exact reasons for these influences are unknown and are frequently unexplained by the models that otherwise mirror the process behavior with reasonable accuracy.

A good example is given in Figure 6–3. A price ceiling was imposed on a double oral auction market such as those discussed in Figure 6–1. In both cases, the ceiling was nonbinding, being at the equilibrium price in the experiment reported in Figure 6–3 and \$0.05 above the equilibrium price for the other experiment. Notice that the prices do not converge tightly to the equilibrium, as is usual; instead, they hang below the equilibrium. Prices will jump immediately to the equilibrium when the ceiling is removed, as demonstrated in Figure 6–4. Price floors exhibit the opposite behavior.

Exactly why these phenomena exist is not clear, and exactly how they should be characterized can be debated, but they are not the only ones that have been discovered. One-sided oral

auctions tend to converge from above (below) when there is an oral bid (offer) (Plott and Smith 1978; Smith 1964). Price posting has an independent effect. If sellers (buyers) post prices, then prices approach equilibrium from above (below). Efficiency for posted price markets is low, and the adjustments of such markets to parameter changes is slow.

Other institutional perturbations have been studied. Vernon Smith and his colleagues at the University of Arizona have been developing and exploring the behavioral properties of computerized markets. Isaac and Plott (1981b, forthcoming) studied institutions that permit the formation of conspiracies. Friedman and Hoggatt (1980) completed a definitive work for the cases of very small competitors interacting over a long period of time.

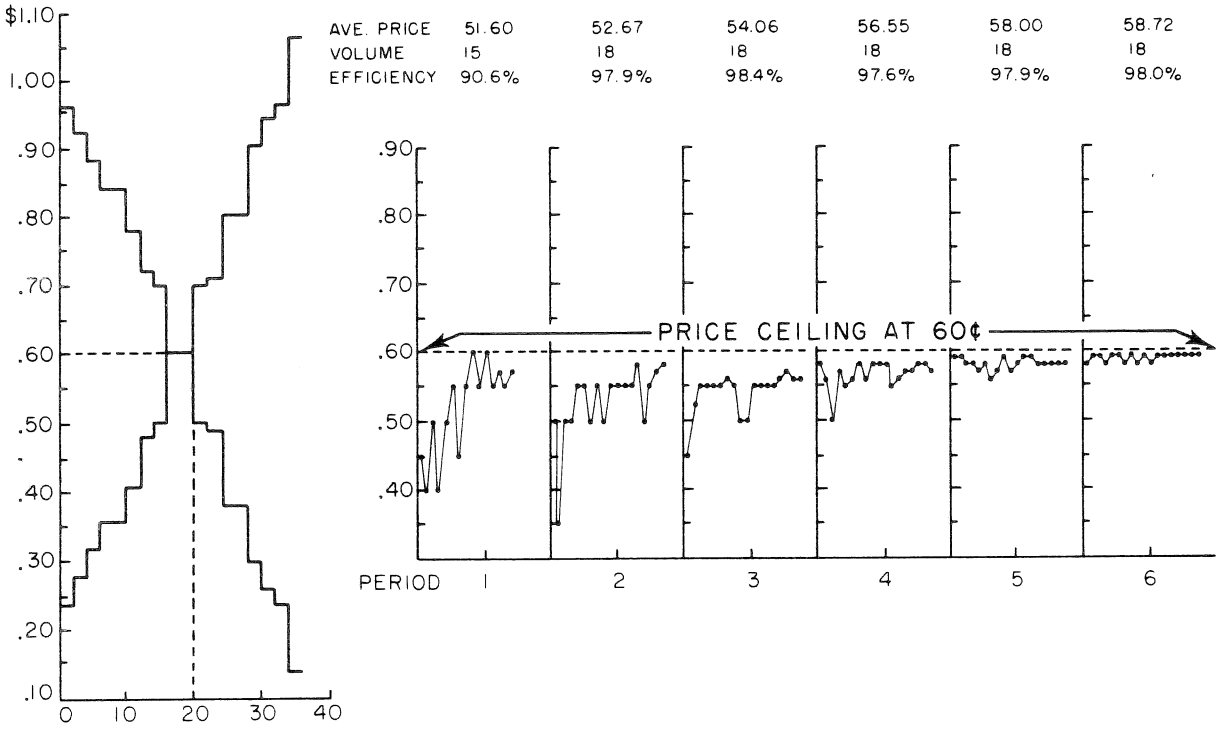
All these institutional variations can exert independent and measurable influences on market behavior. The results from some of these experiments were surprising and are not fully understood theoretically even now. Those interested in theory frequently begin with the simplest cases. The results from simple laboratory environments provide theorists with ample challenges.

The second kind of measurement objectives are those intended to provide econometric measurements. Many economic models have broad policy implications, but the nature and sometimes direction of the implications can be sensitive to the parameters of the models. Several field experiments have been designed to provide information about the parameters of such economic models.

Perhaps the first large-scale field experiment was an attempt to measure the labor supply response that would result from a program of income maintenance. The motivation for this question was to help ascertain the possible consequences of a negative income tax program. An original experiment was conducted in New Jersey and Pennsylvania from 1968 to 1972. This experiment was followed by several others. A summary analysis (Keeley et al. 1978) evaluating the labor response to alternative nationwide national income tax (NIT) programs was published in 1978.

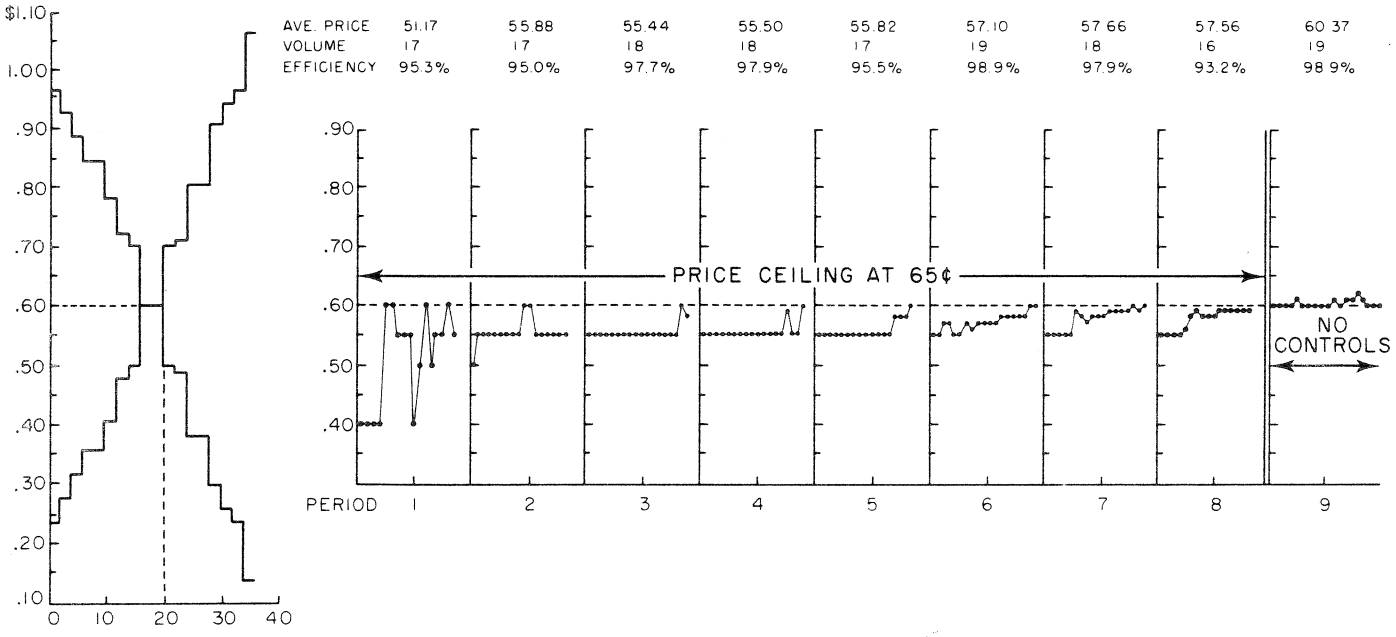
Since the NIT experiments, attempts to measure responsiveness to other types of economic incentive programs have been

Figure 6–3 Market with a nonbinding price ceiling at equilibrium



Source: Isaac, R.M. and C.R. Plott, 1981a, forthcoming. "Price Controls and the Behavior of Auction Markets: An Experimental Examination." *American Economic Review* 71 (June). Reprinted with permission.

Figure 6–4 Market with a nonbinding price ceiling above equilibrium



Source: Isaac, R.M. and C.R. Plott, 1981a, forthcoming. "Price Controls and the Behavior of Auction Markets: An Experimental Examination." *American Economic Review* 71 (June). Reprinted with permission.

initiated. A series of experiments with health insurance have been attempted to measure the extent to which individuals' out-of-pocket payments will affect their use of the health care system and the extent to which differential modes of payment will influence health care.[5] The results indicate considerable sensitivity of demand to the structure of coverage (Newhouse, Phelps, and Schwartz 1974). Measures of health status are also being developed (Brook et al. 1979). It is unlikely that health insurance policies will be developed without relying heavily on the results of these studies.

Several other field experiments exist. Systematic measurements of the responsiveness of housing supply and demand to assistance payments have been initiated with a goal of estimating the likely effects of implementation at the national level.[6] The econometric work in the electricity pricing area has been so successful that the controversies about systematic responsiveness to price no longer exist. The issue has now been refined to one of systematic experimental measurement being funded and explored by the Department of Energy.

In the environmental area, a "bidding game" is being explored as a tool to elicit individual preferences about environmental variables (Brookshire et al. 1979; Rand Corporation 1979). The willingness-to-pay measures are then employed to evaluate environmental conditions and alternatives.

An important observation is in order about experiments designed to provide field measurements. For the most part, those interested in using the results of such field experiments plan to use the measurements as parameters in models that have broad policy implications. This means that the experimental design itself must be developed to reflect the background economic theory. A great deal of theorizing is a necessary component of any such experiment, and the design should reflect this.[7] In addition, contingent strategies must be planned ahead in anticipation of at least some of the many things that can (and do) go wrong.

The technology of designing, monitoring, and controlling a large-scale experiment is not easily mastered. Most papers published on the topic deal with methodological problems that were encountered during a project. Several experiments have terminated with conclusions only about how not to conduct an

experiment. Nevertheless the methodology has produced substantive results, and more are to be expected.

SIMULATION

When most people think about experimentation as applied to policy analysis, they think of simulation. A policy is going to be imposed on a social system. Simulations attempt to recreate a particular social situation on a smaller scale in order to provide decisionmakers with some experience of how such a situation might evolve.

If there is no theory to indicate which variables are unimportant, the complexity of the small situation must mirror the complexity of the large as closely as possible. Furthermore, without theory to unify the observations, the experiment must be conducted enough times to assure the statistical "validity" of any asserted pattern in the results. Thus theory, even in the case of simulation, serves to simplify the experimental process. The more that accepted theory can be invoked, the less the experimental process needs to mirror the natural analog. The tendency of some scholars to reject experimental methods as irrelevant may reflect a basic interest in simulation. Such critics seem to be unaware of the role of theory, but they are aware of the complexities of the situation (and the impossibility of recreating it).

Two attempts to apply new laboratory methods in a policy simulation have been completed. The first involved the (potential) regulation of dry-bulk barge traffic (Hong and Plott, forthcoming). The question was whether barges should post their rates with the Interstate Commerce Commission (ICC) similar to many other rate-posting policies in the transportation industry. Laboratory markets were created that, from an economic theory point of view, were simply scaled down versions of the dry-bulk traffic market. Thus, theory played a key role. Demand statistics, relative size of market participants, cost factors, and demand cycles were equal (except for a scale factor) to the measurements for the 1971 industry. Market performance was studied under the existing mode of organization (a telephone market) and the proposed mode of organization

(price posting). The inefficiency, inflating tendencies, and slow adjustment properties characteristic of posted-price laboratory processes were all replicated. The simple nature of the laboratory process allowed a major problem with price-posting systems to be clearly isolated and understood. And the results, when scaled up to market sizes, at least provided a crude model of what might be expected.

In the second study using laboratory methods to analyze a complex and controversial policy problem (Grether, Isaac, and Plott 1979), the strategy of experimentation was not so much one of simulation as it was one of demonstration. The experiments were used to demonstrate that accepted general theory would in fact work as claimed in laboratory processes that had the prominent features of a more complex process.

Thus in both cases of "simulation" the role played by theory is critical. From a policy advocate's point of view these studies also involve a subtle but important change in the structure of the debate process. Those critical of the studies and related predictions must articulate a reason for the studies to be judged as wrong. With such criticisms in hand, however, the analyst can then examine them with perturbations of the original experiments. One simply conducts additional experiments for which the criticisms do not apply. Presumably at some point sufficient evidence is accumulated to support a decision.

Field demonstration projects are not uncommon and are in a sense simulations, but they generally lack controls. When neither controls nor replications exist, one might claim that the methods are not experimentation. However, the lines of distinction are hard to draw in practice. A program that seems to lie right on the boundary between the two is the Experimental Technology Incentive Program (ETIP) of the National Bureau of Standards. This program of administrative experimentation has involved itself in a wide range of regulatory problems (regulation by the Federal Communications Commission of Western Union's public message service telegrams, requirements by the Environmental Protection Agency for registering pesticides and reducing regulatory lag, employment of postmarketing surveillance of drugs by the Food and Drug Administration). The ETIP explorations are reminiscent of a feedback control device for understanding the consequences of a particular ongoing policy at a particular time. It is an alternative to the traditional command and control strategies that have been used in various forms to effect change in the process of regulation. And, the procedures themselves might be candidates as replacements for proposed rulemaking procedures.

Once a policy has been established, a whole host of institutional reactions are initiated as interested parties who were unknown at the outset reveal themselves through whatever political and bureaucratic means are at their disposal. The researchers at ETIP are exploring methods of monitoring the state of policies, formulating feedback systems, recommending modifications of policies as vested interests and unplanned consequences reveal themselves and documenting the experiences for those who are considering implementation of related policies.

Other explorations with alternative forms of administrative structure exist. School voucher experiments involve studies of complicated administrative structures (Levinson 1976). An administrative agency experiment was incorporated as part of the Experimental Housing Allowance Program. The idea involved the use of different types of agencies in the administering of a program. Generalizations from these experiences are difficult because the small number of observations and large number of variables admit the possibility of many competing hypotheses to explain what was observed. Naturally, to the extent that controls exist, the number of competing hypotheses are reduced.

PROCESS DESIGN

The final objective of experimental methods is in a sense the opposite of the other objectives. Experimentation, and especially laboratory experimentation, can be used to provide experience with completely new modes of organization that have no "natural" origin. Such "synthetic" processes cannot be studied historically since none have ever existed.

With the previously discussed objectives, an ongoing process is chosen, and the ultimate objective is to understand how and

why it (or slight variations of it) functions the way it does. Technically put, the objectives are sometimes to find a model (or mechanism) that reflects the essence of a process and will help predict the consequences of slight changes in the process. The design objectives are essentially the reverse. Designs are motivated by a mechanism (a mathematical model, a body of theory) that is perhaps completely devoid of operational detail. The task is to find a system of institutions—the rules for individual expression, information transmittal, and social choice—a "process" that mirrors the behavioral features of the mechanism. The theory suggests the existence of processes that perform in certain (desirable) ways, and the task is to find them. This is a pure form of institutional engineering.

Several examples exist in the laboratory experimental literature. Using abstract theory as a guide, Plott and Levine (1978) designed a class of agenda that are consistent with many aspects of *Robert's Rules of Order* but can be used to influence if not determine the decisions of voting committees. Another example is a series of papers that were stimulated by the literature on abstract adjustment processes found in mathematical economics. Economists, while exploring the mathematical structure of a class of abstractly conceived economic processes, discovered that theoretical processes exist which solve the classical public goods problem. This problem constitutes the core of theory about market failure where externalities, pollution, or public goods are present.[8] Vernon Smith (1978, 1979a, 1979b), using the mathematical models as a guide, designed a series of decisionmaking systems in operational terms and used laboratory techniques to explore their behavioral properties.

Stimulated by the possibility of applying laboratory experimental techniques, Ferejohn, Forsythe, and Noll (1979) and Ferejohn and Noll (1976) examined a problem faced by the Public Broadcasting Service (PBS). Independent stations face a periodic problem of selecting programs to be jointly programmed and financed. Both the particular programs and cost sharing must be determined within a single process. It is the classic problem of joint provision of a public good. The project undertaken by Ferejohn, Forsythe, and Noll involved tailoring processes suggested in the literature to meet the needs of the

PBS system. The properties and problems with these alternative systems are being evaluated by laboratory experimental techniques prior to the design of any field experiments or implementation.

Some of the field experimentation and demonstration projects discussed might be described more accurately as having process design objectives. The voucher experiments, for example, were clearly motivated by a body of theory about how a system of competition for students would work (Weiler 1974). The theory suggests that certain new systems will have particular properties; the experiment is designed to see if they do.

SOME ADVANTAGES OF EXPERIMENTAL METHODS

Experimental methods have some advantages over other research methodologies. They provide an inexpensive means of obtaining experiences with the behavior of processes and the accuracy of models of those processes. Compared with historical techniques and econometric techniques, the expenses are low and the quality of the data is excellent.

In addition, experimental studies tend to be conceived and executed quickly. Even the most complicated of social experiments last only a few years. Laboratory studies, on the other hand, can be completed in a matter of months, if not weeks.

Experimental methods allow a type of flexibility in situations that history frequently fails to provide. Situations can be created for which competing models give substantially different predictions. Thus the opportunity is presented for using an empirical basis for pruning the theoretical tree of the less successful models.

Experimentation provides a type of practical advantage that might at first be overlooked but is extremely important. In the experimental world, everything must be operationalized. Abstract terms (preferences, rights, strategies, information) that frequently serve as parameters on models must be specified operationally. For theorists, this task is sometimes shockingly difficult and leads to the discovery of real ambiguities in the theory. The necessity for operationalization also tends to flush

out policy bugs that could spell potential failure for a politically sensitive course of action. (A record form may be ambiguous. Instructions may be difficult to understand. Information that is obvious to policymakers may be too obscure to participants.) The small, unanticipated problems that can result from necessarily incomplete planning and can cost thousands of dollars if discovered too late tend to show up in the experiment. The only cost at the experimental stage is that a relatively inexpensive experiment must be scrapped and redone.

Some of the primary advantages of experimental methods are derived from the possibility of replication. Those skeptical of the results or those who do not trust the processors of the data can conduct the experiment themselves. In the physical sciences, replication of the results of others has developed as a standard and legitimate academic endeavor. Most researchers involved with laboratory methods in economics and political science replicate the results of others before relying on them in their own research. Those of us involved in experimental methods are always amazed by the variety of results that really do replicate in spite of the differences in subject pools, experimental settings, experimenters, and in some cases, experimental techniques.

Not only is replication the vehicle through which experience with particular phenomena become shared, it is the process through which the competition among explanations gets resolved. Almost any feature of an experiment or unplanned event during an experiment can provide the basis for an explanation of all the empirical results. Here are some typical examples: people did not understand the instructions; the subject pool was special; the number of participants or some important features of their lives (such as alternative job opportunities in a labor supply experiment) were not controlled; the incentives were not sufficiently large; the incentives were too large; the experimenter intervened to produce a "Milgrom" effect or the results are due to a "Hawthorn" effect.

Such sources of competing hypotheses only become resolved through replication and modification of the experimental design. Until the experiment can be performed again with controls for the hypothesized effects absent, the criticisms remain. Of course, a decisionmaker may decide that the evidence is sufficient to justify an action in spite of the existing competing theories, but that is a matter of judgment. The nagging scientific questions necessarily remain until the alternative experiments are conducted. Controversial claims may take many experimental series to resolve and, for large-scale social experiments, may be expensive. The point is that the conflicts are in principle resolvable and that the method of resolution frequently involves more experiments. So, those thinking about the use of experimental methods as research tools should generally think in terms of a series (perhaps small) of experiments rather than only one experiment.

It seems that one can learn nothing conclusive from experimental methods about how naturally occurring processes will respond to various policies. Even in the physical sciences, experimental methods have not developed to that exalted state. It is clear, however, that one can learn much from experimentation about models that purport to predict what will happen. And, one can get considerable experience with the operations of simple versions of the process. In other words experimentation can lead to informed guesses and, as our ability to properly pose questions improves, the methods will probably be applied with increasing frequency in many phases of regulatory research.

NOTES

1. Similarities and divergencies in methodologies have not been systematically explored. Some excellent discussions can be found dealing with the methodologies that have evolved from psychologically oriented studies (Campbell and Stanley 1963; Cook and Campbell 1976; McGrath 1964; Weick 1967), but currently none exist that deal with or attempt to summarize those which have evolved from an economics or mathematical political science orientation. The only attempts are Kagel and Battalio (1980), Plott (1979), and Smith (1980).
2. Psychologists mustering a defense against such claims have begun to document a "knew-it-all-along" effect (Fischhoff 1975, 1977; Fischhoff and Beyth 1975; Wood 1978). While the research purposes were not an attempt to muster a defense against such attacks the results certainly aid in that capacity.

3. Consumer plus producer surplus is maximized here if and only if the group maximizes the total possible earnings from the experiment. This efficiency measure, first developed by Plott and Smith (1978), is now being used as a tool to evaluate the performance of competing processes within laboratory experiments.
4. It is interesting to note that these limitations to the economic theory of preference as applied to humans are being discovered by psychologists at the same time that the economic theory of preference is beginning to evolve as a general theory explaining psychological experiments involving the choice behavior of nonhumans (Kagel et al. 1975).
5. For a summary listing of objectives, see Newhouse (1979).
6. Ferber and Hirsch (1978) include a review. See also Carlson and Heinberg (1978) and Rand Corporation (1979).
7. An excellent collection of papers on this topic can be found in Aigner and Morris (1979). See also Barnett and Lowry (1979) for a summary of results.
8. For a review of the nature of the discovery, see *Public Choice* (1977).

REFERENCES

Aigner, D.J., and C.N. Morris, eds. 1979. "Experimental Design in Econometrics." *Journal of Econometrics* (special supplement) 11, no. 1 (September) 1–205.

Barnett, C. Lance, and Ira S. Lowry. 1979. "How Housing Allowances Affect Housing Prices." R-2452-HUD, Rand Corporation, Santa Monica, Calif.

Belovicz, M.W., 1979. "Sealed-Bid Auctions: Experimental Results and Applications." In *Research in Experimental Economics,* vol. 1, edited by Vernon L. Smith. Greenwich, Conn.: JAI Press.

Berl, J.; R.D. McKelvey; P.C. Ordeshook; and Mark Winer. 1976. "An Experimental Test of the Core in a Simple N-Person Cooperative Nonsidepayment Game." *Journal of Conflict Resolution* 20 (September):453–79.

Brook, R.H.; J.E. Ware, Jr.; A. Davies-Avery; A.L. Stewart; C.A. Donald; W.H. Rogers; K.N. Williams; and S.A. Johnson. 1979. "Overview of Adult Health Status Measures Fielded in Rand's Health Insurance Study." *Medical Care* (supplement) 17, no. 7. (July): 1–131.

Brookshire, D.S.; R.C. d'Arge; W.D. Schulze; and M.A. Thayer; 1979. *Methods Development for Assessing Tradeoffs in Environmental Management,* vol. 2: Experiments in Valuing Non-Market Goods: A Case Study of Alternative Benefit Measures of Air Pollution Control in the South Coast Air Basin of Southern California. EPA-600/6-79-001b. Washington, D.C., February.

Cain, Glen G., and Harold W. Watts, eds. 1973. *Income Maintenance and Labor Supply: Econometric Studies.* Chicago: Rand McNally.

Campbell, D.T., and J.C. Stanley. 1963. *Experimental and Quasi-Experimental Designs for Research.* Chicago: Rand McNally.

Carlson, D.B., and J.D. Heinberg. 1978. "How Housing Allowances Work: Integrated Findings from the Experimental Housing Allowance Program." Washington, D.C.: Urban Institute.

Conlisk, John, and Harold W. Watts. 1969. "A Model for Optimizing Experimental Design for Estimating Response Surfaces." In *Proceedings of the Social Statistics Section, 1969,* pp. 150–56. Washington, D.C.: American Statistical Association.

Cook, T.D., and D.T. Campbell. 1976. "The Design and Conduct of Quasi-Experiments and True Experiments in Field Settings." In *Handbook of Industrial and Organizational Psychology,* edited by M.D. Dunnette, pp. 224–46. Chicago: Rand McNally.

Ferber, R., and W.Z. Hirsch. 1978. "Social Experimentation and Economic Policy: A Survey." *Journal of Economic Literature* 16 (December): 1379–1414.

Ferejohn, J.A.; R. Forsythe; and R. Noll. 1979. "An Experimental Analysis of Decisionmaking Procedures for Discrete Public Goods." In *Research in Experimental Economics,* vol. 1, edited by Vernon L. Smith, pp. 1–58. Greenwich, Conn.: JAI Press.

Ferejohn, J.A., and R. Noll. 1976. "An Experimental Market for Public Goods: The PBS Station Program Cooperative." *American Economic Review* 66 (May): 267–73.

Fiorina, M.P., and C.R. Plott. 1978. "Committee Decisions under Majority Rule: An Experimental Study." *American Political Science Review* 72 (June):575–98.

Fischhoff, B. 1975. "Hindsight = Foresight: The Effect of Outcome Knowledge on Judgment under Uncertainty." *Journal of Experimental Psychology: Human Perception and Performance* 1 (August):288–99.

Fischhoff, B. 1977. "Perceived Informativeness of Facts." *Journal of Experimental Psychology: Human Perception and Performance* 3 (May):349–58.

Fischhoff, B., and R. Beyth. 1975. "'I Knew It Would Happen'—Remembered Probabilities of Once-Future Things." *Organizational Behavior and Human Performance* 13:1–16.

Friedman, J.W., and A.C. Hoggatt. 1980. *An Experiment in Noncooperative Oligopoly. Research in Experimental Economics, Supplement 1.* Greenwich, Conn.: JAI Press.

Grether, David M. 1980. "Bayes Rule as a Descriptive Model: The Representativeness Heuristic." *Quarterly Journal of Economics* 95 (November):537–57.

Grether, D.M.; R.M. Isaac; and C.R. Plott. 1979. "Alternative Methods of Allocating Airport Slots: Performance and Evaluation." CAB79-C-73, Polinomics Research Laboratories, P.O. 5932, Pasadena, Calif. 91107.

Grether, David M., and Charles R. Plott. 1979. "Economic Theory and the Preference Reversal Phenomenon." *American Economic Review* 69 (September):623–38.

Hershey, J.C., and P.J.H. Schoemaker. 1980. "Risk Taking and Problem Context in the Domain of Losses: An Expected Utility Analysis." *Journal of Risk and Insurance.* Vol. 47, No. 1: 111–132.

Hong, J., and C.R. Plott. Forthcoming. "Implications of Rate Filing for Domestic Dry Bulk Transportation on Inland Waters: An Experimental Approach." Forthcoming. *Bell Journal of Economics.*

Isaac, R.M., and C.R. Plott. 1981a. Forthcoming. "Price Controls and the Behavior of Auction Markets: An Experimental Examination." *American Economic Review* 71 (June).

Isaac, R.M., and C.R. Plott. 1981b. Forthcoming. "The Opportunity for Conspiracy in Restraint of Trade: An Experimental Study." *Journal of Economic Behaviour and Organization.*

Kagel, John H., and R.C. Battalio. 1980. "Token Economy and Animal Models for the Experimental Analysis of Economic Behavior." In *Evaluation of Econometric Models,* edited by J. Kmenta and J. Ramsey. New York: Academic Press.

Kagel, John H.; R.C. Battalio; H. Rachlin; L. Green; R.L. Basmann; and W.R. Klemm. 1975. "Experimental Studies of Consumer Demand Behavior Using Laboratory Animals." *Economic Inquiry* 13:22–38.

Kahneman, D., and A. Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47 (March):263–91.

Keeley, M.C.; P.K. Robins; R.G. Spiegelman; and R.W. West. 1978. "The Estimation of Labor Supply Models Using Experimental Data." *American Economic Review* 68 (December):873–87.

Kershaw, David, and Jerilyn Fair. 1976. *The New Jersey Income-Maintenance Experiment,* vol. 1: *Operations, Surveys and Administration.* New York: Academic Press.

Kunreuther, H.; R. Ginsberg; L. Miller; P. Sagi; P. Slovic; B. Borkan; and N. Katz. 1978. *Disaster Insurance Protection: Public Policy Lessons.* New York: Wiley.

Kunreuther, H.; J. Lepore; L. Miller; J. Vinso; J. Wilson; B. Borkan; B. Duffy; and N. Katz. 1978. *An Interactive Modeling System for Disaster Policy Analysis.* Boulder, Colo.: Institute of Behavioral Science.

Laing, J.D., and S. Olmsted. 1978. "An Experimental and Game Theoretic Study of Committees." In *Game Theory and Political Science,* edited by P.C. Ordeshook. New York: New York University Press.

Levinson, Eliot. 1976. "The Alum Rock Voucher Demonstration: Three Years of Implementation." P-5631. Santa Monica, Calif.: Rand Corporation, April.

Lichtenstein, S., and P. Slovic. 1971. "Reversals of Preference between Bids and Choices in Gambling Decisions." *Journal of Experimental Psychology* 89 (July):46–55.

Lichtenstein, S., and P. Slovic. 1973. "Response-Induced Reversals of Preference in Gambling: An Extended Replication in Las Vegas." *Journal of Experimental Psychology* 101 (November):16–20.

Lindman, H.R. 1971. "Inconsistent Preferences among Gambles." *Journal of Experimental Psychology* 89 (August):390–97.

Lowry, Ira S. 1977. *An Overview of the Housing Assistance Supply Experiment.* P-5976 Santa Monica, Calif.: Rand Corporation.

McGrath, J.E. 1964. "Toward a 'Theory of Method' for Research on Organizations." In *New Perspectives in Organizational Research,* edited by W.W. Cooper, H.L. Leavitt, and M.W. Shelby, pp. 533–57. New York: Wiley.

McKelvey, R.D., and P.C. Ordeshook. 1978. "An Experimental Test of Several Theories of Committee Decision Making under Majority Rule." Pittsburgh: Carnegie-Mellon University. April. Mimeo.

McKelvey, R.D.; P.C. Ordeshook; and M.D. Winer. 1978. "The Competitive Solution for N-Person Games without Transferable Utility, with an Application to Committee Games." *American Political Science Review* 72 (June):599–615.

Mitchell, Bridger M.; Willard G. Manning, Jr.; and Jan Paul Acton. 1978. *Peak-Load Pricing: European Lessons for U.S. Energy Policy.* Cambridge, Mass.: Ballinger Publishing Company.

Newhouse, J.P. 1979. "A Summary of the Experimental Portion of the Rand Health Insurance Study." Santa Monica, Calif.: Rand Corporation.

Newhouse, J.P.; C.E. Phelps; and W.B. Schwartz. 1974. "Policy Options and the Impact of National Health Insurance." R-1528-

HEW/OEO. Santa Monica, Calif.: Rand Corporation, June: *New England Journal of Medicine*, 290 (June):1345–59.

Palmer, John L., and Joseph A. Pechman, eds. 1978. *Welfare in Rural Areas: The North Carolina-Iowa Income Maintenance Experiment*. Washington, D.C.: The Brookings Institution.

Pechman, Joseph A., and P. Michael Timpane, eds. 1975. *Work Incentives and Income Guarantees: The New Jersey Negative Income Tax Experiment*. Washington, D.C.: The Brookings Institution.

Plott, Charles R. 1979. "The Application of Laboratory Experimental Methods to Public Choice." In *Collective Decision Making: Applications from Public Choice Theory*, edited by Clifford S. Russell. Washington, D.C.: Resources for the Future.

Plott, C.R., and W.P. Rogerson. 1979. "Committee Decisions under Majority Rule: Dynamic Theories and Experimental Results." Social Science Working Paper no. 280. Pasadena: California Institute of Technology, August.

Plott, C.R., and M.E. Levine. 1978. "A Model of Agenda Influence on Committee Decisions." *American Economic Review* 68 (March):146–60.

Plott, C.R., and V.L. Smith. 1978. "An Experimental Examination of Two Exchange Institutions." *Review of Economic Studies* 45 (February):133–53.

Public Choice 29, special supplement to Spring 1977. Edited by T. Nicolaus Tideman.

Randall, A.; B. Ives; and E. Eastman. 1974. "Bidding Games for Valuation of Aesthetic Environmental Improvements." *Journal of Environmental Economics and Management* 1:132–49.

Rand Corporation. 1979. *Fifth Annual Report of the Housing Assistance Supply Experiment*. R2434/HUD (June) Santa Monica, Calif.

Rapoport, Amnon; J.P. Kahan; and T.S. Wallsten. 1976. "Sources of Power in Four-Person Apex Games." L.L. Thurstone Psychometric Laboratory Report no. 151. Chapel Hill: University of North Carolina, October.

Savage, L.J. 1954. *Foundations of Statistics*. New York: Wiley.

Schoemaker, P.J.H., and H. Kunreuther. 1979. "An Experimental Study of Insurance Decisions." *Journal of Risk and Insurance*. vol. 46 no.4:603–18.

Slovic, Paul. 1975, "Choice between Equally Valued Alternatives." *Journal of Experimental Psychology: Human Perception and Performance* 1 (August):280–87.

Smith, Vernon L. 1964. "The Effect of Market Organization on Competitive Prices." *Quarterly Journal of Economics* 78 (May):181–201.

Smith, Vernon L. 1967. "Experimental Studies of Discrimination Versus Competition in Sealed-Bid Auction Markets." *Journal of Business* 40 (January):56–84.

Smith, Vernon L. 1976a. "Bidding and Auctioning Institutions: Experimental Results." In *Bidding and Auctioning for Procurement and Allocation*, edited by Yakov Amihud. New York: New York University Press.

Smith, Vernon L. 1976b. "Experimental Economics: Induced Value Theory." *American Economic Review* 66 (May):274–79.

Smith, Vernon L. 1978. "Experimental Mechanisms for Public Choice." In *Game Theory and Political Science*, edited by P. C. Ordeshook, pp. 323–55. New York: New York University Press.

Smith, Vernon L. 1979a. "Incentive Compatible Experimental Processes for the Provision of Public Goods." In *Research in Experimental Economics*, vol. 1, pp. 59–168. Greenwich, Conn.: JAI Press.

Smith, Vernon L. 1979b. "An Experimental Comparison of Three Public Good Decision Mechanisms." *Scandinavian Journal of Economics* 81:198–215.

Smith, Vernon L. 1980. "Relevance of Laboratory Experiments to Testing Resource Allocation Theory." In *Evaluation of Econometric Models*, edited by J. Kmenta and J. Ramsey. New York: Academic Press.

Tsao, Che, and Anthony Vignola. 1977. "Price Discrimination and the Demand for Treasury's Long Term Securities." Preliminary report circulated privately.

Tversky, A. 1969. "Intransitivity of Preferences." *Psychological Review* 76 (January):31–48.

Watts, Harold W., and Albert Rees, eds. 1977. *The New Jersey Income-Maintenance Experiment*, vol. 2: *Labor-Supply Responses*. New York: Academic Press.

Watts, Harold W., and Albert Rees, eds. 1977. *The New Jersey Income-Maintenance Experiment*, vol. 3: *Expenditures, Health and Social Behavior; and the Quality of the Evidence*. New York: Academic Press.

Weick, K.E. 1967. "Organizations in the Laboratory." In *Methods of Organizational Research*. Pittsburgh: University of Pittsburgh Press.

Weiler, D. 1974. "A Public School Voucher Demonstration: The First Year at Alum Rock." R-1495-NIE. Santa Monica, Calif.: Rand Corporation, June.

Wood, G. 1978. "The Knew-It-All-Along Effect." *Journal of Experimental Psychology: Human Perception and Performance* 4, no. 2:345–53.